# Health Insurance

## cost

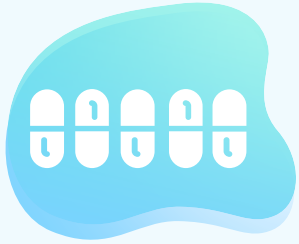
# METHODOLOGY

INTRODUCTION

EDA

Feature engineering

Linear regression models



CONCLUSION


01
INTRODUCTION

Health Insurance cost

# INTRODUCTION

Today we'll explore a Data Set dedicated to the price of treatment of various patients. the price of treatment depends on several factors: designation, form of clinic, town of residence, age then on. we've got no data on the diagnosis of patients. However, we have alternative information which will facilitate U.S.A. to create a conclusion regarding the health of patients and observe regression analysis.

# FEATURES

## SEX

insurance contractor gender, female, male

## AGE

age of primary beneficiary

## BMI

Body mass index , providing an understanding of body.

## CHILDREN

Number of children covered by health insurance

## SMOKER

Smoking , yes or No

## REGION

the beneficiary's residential area in the US, south , north

## CHARGES

Individual medical costs billed by health insurance

predict insurance costs?

# Our target Charges

# 02

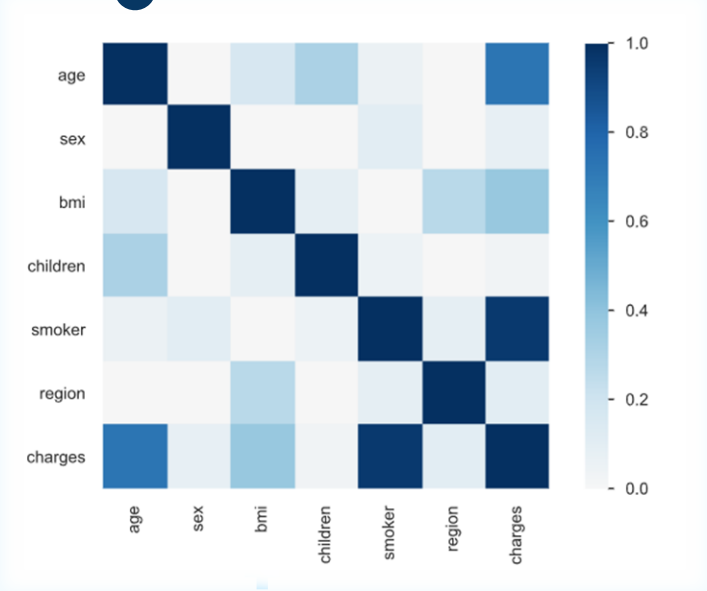
# EDA

Free Null 01

Correlations 03

02 No Duplicate

# Correlations



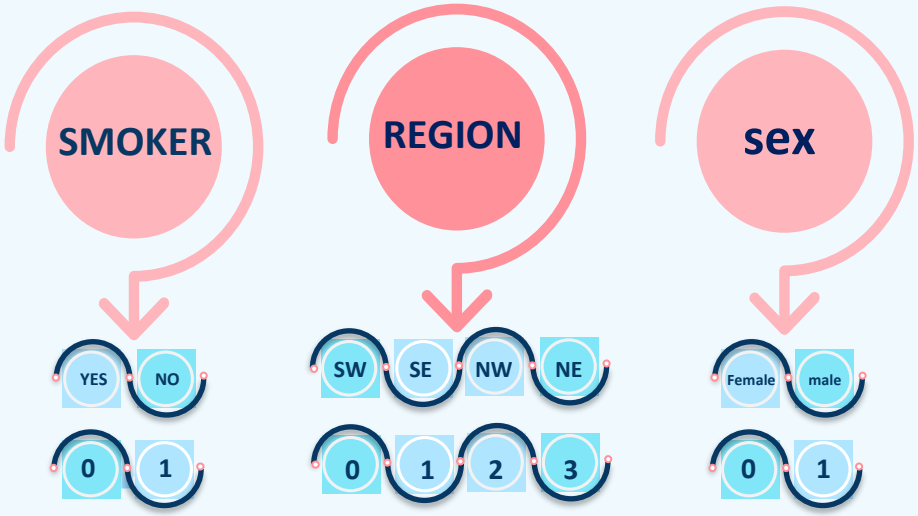
# Correlations Charges Number

| Columns | Correlation |
|---|---|
| Region | - 0.0062 |
| SEX | 0.0572 |
| Children | 0.0679 |
| BMI | 0.1983 |
| AGE | 0.2990 |
| Smoker | 0.7872 |
| Charges | 1.0000 |

# Charges and Smokers..


03
Feature
Engineering

In our project we used Label Encoder  which converts from non-numerical labels to numerical.




04
Linear Regression
Models

# MODELS SCORES

## Linear Regression

Train : 0.749008
validation : 0.742403

## Random Forest

Train : 0.974000
validation : 0.869000

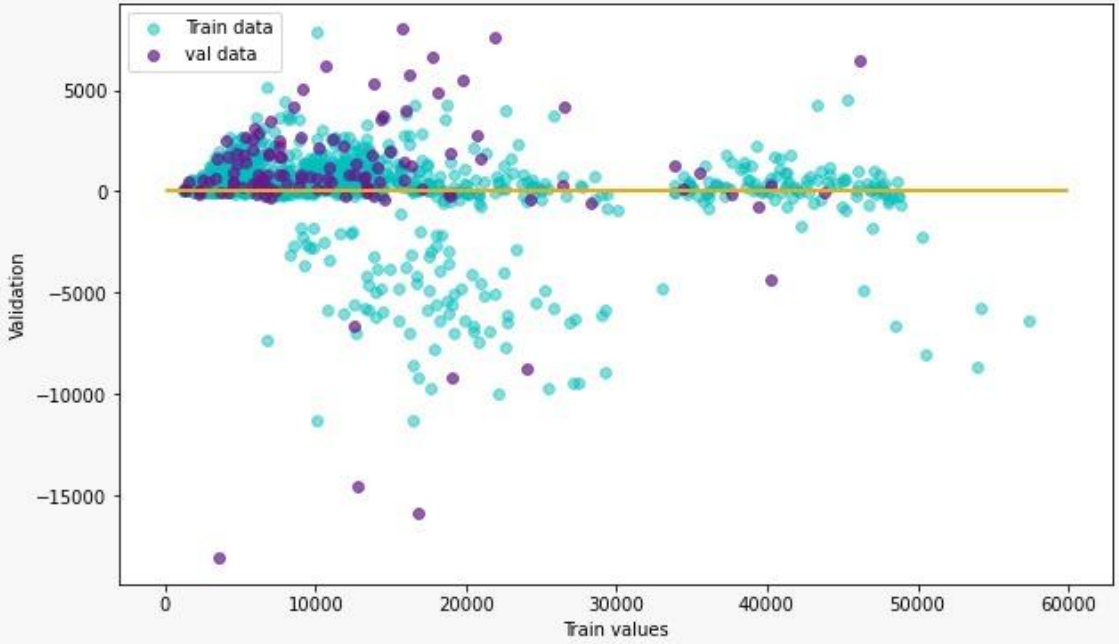
## Polynomial Features

Train : 0.843331
validation : 0.822541

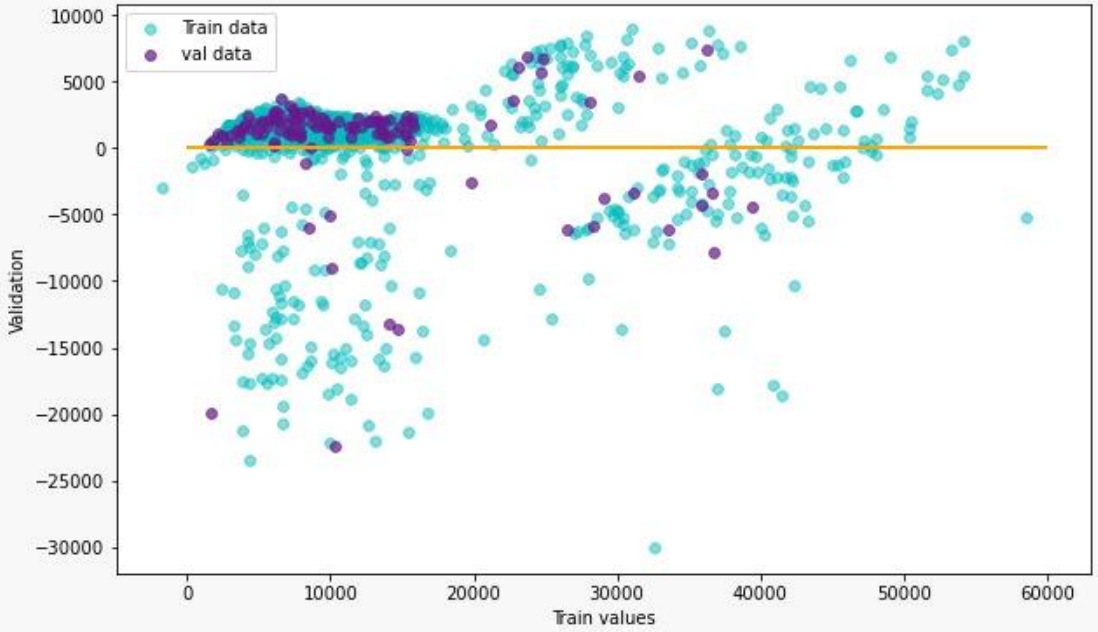
## Lasso

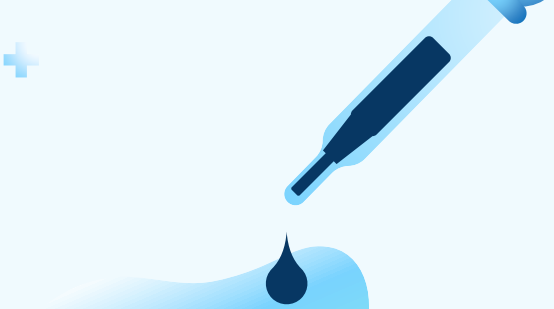
Train : 0.749008
validation : 0.742103

# Random Forest Model



# Polynomial Model





# 05

# CONCLUSION

# CONCLUSION

- We see a strong relationship between smokers and the Charges of insurance.

- In the end, the Polynomial has been selected because it's the highest ratio experiment (Best Fit).

# THANKS

For watching ..

Health Insurance cost


**Presented by**

Rawan Alshehri
Khalid Alharbi
Abdulrhman Alqurashi